

SPEECH INPUT DEVICE

(2)

Patent Number: JP10312193
Publication date: 1998-11-24
Inventor(s): ONO TAKESHI;; KISHI NORIMASA
Applicant(s): NISSAN MOTOR CO LTD
Requested Patent: ☐ JP10312193
Application Number: JP19970135793 19970512
Priority Number(s):
IPC Classification: G10L3/00; G10L3/00; G10L3/00; G08G1/0969
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To provide a voice input device capable of inputting a large vocabulary speech with the number of less input times.

SOLUTION: A microphone 500 and an A/D converter 501 fetch speech, and convert it to a digital speech signal to output it to a signal processor 5. A hierarchized word dictionary is stored in an external storage 503. A monitor 60 with a touch panel displays a recognition result, and a partial dictionary is revised by a touch input. The signal processor 5 inputs the partial dictionary of the prescribed lowest rank hierarchy first, and operates a matching degree between the voice signal and a word in the dictionary, and displays the word with the highest matching degree on the monitor 60 while containing its high- order hierarchical structure. When its word isn't the word to be recognized, the hierarchy to be revised is touch inputted. Its touch signal is detected, and the corresponding dictionary is read in from the external storage 530. Thereafter, the hierarchy is decided by the speech, and the lowermost layer dictionary under its hierarchy is read in. An object is uttered newly, and is voice recognized by collating with the dictionary.

Data supplied from the esp@cenet database - I2

(書誌+要約+請求の範囲)

- (19)【発行国】日本国特許庁(JP)
(12)【公報種別】公開特許公報(A)
(11)【公開番号】特開平10-312193
(43)【公開日】平成10年(1998)11月24日
(54)【発明の名称】音声入力装置
(51)【国際特許分類第6版】

G10L 3/00 571
551
561

G08G 1/0969

【FI】

G10L 3/00 571 G
551 J
561 C

G08G 1/0969

【審査請求】未請求

【請求項の数】10

【出願形態】FD

【全頁数】13

- (21)【出願番号】特願平9-135793
(22)【出願日】平成9年(1997)5月12日
(71)【出願人】

【識別番号】000003997

【氏名又は名称】日産自動車株式会社

【住所又は居所】神奈川県横浜市神奈川区宝町2番地

(72)【発明者】

【氏名】大野 健

【住所又は居所】神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内

(72)【発明者】

【氏名】岸 則政

【住所又は居所】神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内

(74)【代理人】

【弁理士】

【氏名又は名称】菊谷 公男 (外3名)

(57)【要約】

【課題】 少ない入力回数で、大語彙音声入力できる音声入力装置を提供する。

【解決手段】 マイク500とA/Dコンバータ501は音声を取り込んでデジタルの音声信号に変えて信号処理装置5に出力する。外部記憶装置503には階層化した単語辞書を記憶させている。タッチパネル付きモニタ60は、認識結果を表示するとともに、タッチ入力で部分辞書の変更が行なえる。信号処理装置5では、最初に所定の最下位階層の部分辞書を入力し、音声信号と辞書内の単語との一致度を演算し、一致度の最も高い単語をその上位階層構造を含めモニタにおいて表示する。その単語が認識したい単語でなければ、変更したい階層がタッチ入力される。そのタッチ信号が検出され対応した辞書を外部記憶装置503から読み込まれる。その後音声により階層が決定され、その階層下の最下層辞書が読み込まれる。目的語を新たに発話して辞書との照合により音声認識される。

【特許請求の範囲】

【請求項1】 発話音声を情報化して入力する音声入力手段と、複数の単語を含む部分辞書からなるとともに階層構造を持った単語辞書と、使用する部分辞書を決定する部分辞書決定手段と前記音声入力手段からの音声情報と決定された部分辞書内の単語との一致度を演算する演算手段と、演算された単語のうち最も一致度の高い単語を選択して出力する単語選択手段とを有し、前記部分辞書決定手段は、初めに所定の最下層の部分辞書を決定し、階層変更指令を受けた場合に階層変更をして部分辞書を新たに決定することを特徴とする音声入力装置。

【請求項2】 前記部分辞書決定手段は、上位階層への変更指示を受けた場合、指示された上位階層の部分辞書を決定し、入力により前記単語選択手段が選択した単語が示す階層に変更し、該階層下の所定の最下層部分辞書を決定することを特徴とする請求項1記載の音声入力装置。

【請求項3】 前記部分辞書決定手段により初めに決定される部分辞書は前回使用された最下層の部分辞書であることを特徴とする請求項1または2記載の音声入力装置。

【請求項4】 前記部分辞書決定手段により初めに決定される部分辞書は使用頻度の高い最下層の部分辞書であることを特徴とする請求項1または2記載の音声入力装置。

【請求項5】 前記音声入力装置には表示手段が接続され、前記単語選択手段が選択した単語及び上位の階層構造を表示するようにしたことを特徴とする請求項1または2記載の音声入力装置。

【請求項 6】 前記表示手段はタッチパネルを合わせ持ち、使用者によってタッチされた階層に対応した階層変更指令が前記部分辞書決定手段に出力されることを特徴とする請求項 5 記載の音声入力装置。

【請求項 7】 前記表示手段はタッチ入力があったときに、変更可能な階層を表示することを特徴とする請求項 6 記載の音声入力装置。

【請求項 8】 前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、上位階層への変更指示を受けたとし、設定されている最下層部分辞書より上位の全ての部分辞書を決定し、単語選択手段が選択した単語が示す階層に階層変更を行なうことを特徴とする請求項 1 または 2 記載の音声入力装置。

【請求項 9】 前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、設定されている最下層部分辞書及びそれより上位の全ての部分辞書を決定し、単語選択手段が選択した単語が上位部分辞書内の単語であれば、上位階層への変更指示を受けたとし、前記単語が示す階層に階層変更を行なうことを特徴とする請求項 1 または 2 記載の音声入力装置。

【請求項 10】 前記音声入力装置は最初に入力された音声信号を記憶する記憶手段を合わせ持ち、階層が変更されたとき、前記単語選択手段が前記記憶手段に記憶された音声信号を用いて、決定された最下層の部分辞書内の単語との一致度を演算し、出力することを特徴とする請求項 1、2、8 または 9 記載の音声入力装置。

分野

【発明の属する技術分野】この発明は、少ない入力回数で大語彙認識を行なうことができる音声入力装置に関する。

技術

【従来の技術】 音声を識別して情報を読み取る音声入力装置は、手を介さずに指令やデータなどの入力ができるため、障害者用ワープロ、車載ナビゲーションといったような手操作が困難な場面に採用されつつある。典型的な音声入力装置は、図 13 に示すように、音声を取り込み、電気信号に変えて音声信号を出力するマイク 500 と、アナログの音声信号を A/D 変換してデジタル情報に変換させる A/D コンバータ 501 と、音声信号から単語認識をするのに用いられる単語辞書を記憶した外部記憶装置 503 と、CPU とメモリを持ち単語辞書をメモリに読み込ませたうえで CPU が音声信号と辞書内の単語との一致度判定により音声認識を行なう信号処理装置 502 とを有する。スイッチ 505 は装置使用者に操作され音声信号の取り込みタイミングを信号処理装置 502 に与える。モニタ 504 は認識された単語を表示する。

【0003】 一般に音声による入力は、認識対象の語数が多くなる程認識率が低下するため、大語彙認識の場合階層化した辞書が使用される。入力の際、辞書の分類に従って複数

回音声入力を繰り返して階層を進み、認識対象語数を絞り込んだ状態で認識し、誤認識率を低下させるようにしている。

【0004】上記従来の装置における音声認識の詳細について、図14のフローチャートに従って説明する。ここでは、JRの駅名である「桜木町」を音声で入力する場合を示す。このような駅名入力は、自動車のナビゲーション装置の目的地設定などに用いられている。外部記憶装置503には例えば図2のような階層化した単語辞書を記憶させてある。

【0005】まず、ステップ521において、音声入力の初期状態で、信号処理装置502は階層化された単語辞書の、階層の最上位の部分辞書を最初の認識対象とし設定する。ここでは図2中の部分辞書aが設定される。次に、装置使用者はスイッチ505を操作してマイク500に「施設」を発話する。これは「桜木町」が「施設」という範疇に含まれることを使用者が認識しているためである。スイッチ505が操作されたことは、ステップ522において、検出される。これによって音声認識処理が開始される。

【0006】ステップ523では、信号処理装置502内のCPUが設定された階層の部分辞書を外部記憶装置503からメモリに読み込ませる。初期では部分辞書aが読み込まれる。信号処理装置502は、スイッチ505が操作されるまでの音声の平均パワーを演算しており、スイッチ505が操作された後、音声の平均パワーに比べて音声信号の瞬間パワーが所定値以上大きくなったとき、使用者が発話開始したと判断し、ステップ524において音声取り込みを開始する。

【0007】ステップ525では、CPUが入力された音声とメモリに読み込んだ部分辞書内の単語との一致度を演算する。初期では「住所」、「施設」それぞれとの一致度を演算する。

【0008】そして、音声信号の瞬間パワーが所定値以下になった時、使用者の発話が終了したと判断し、ステップ526において、音声の取り込みを終了する。入力音声と部分辞書内の単語との一致度の演算を終了すると、ステップ527において、CPUが最も一致度高い単語を選択する。ここでは「施設」の方が一致度が高いので選択される。

【0009】ステップ528では、選択された一致度の高い単語が認識の結果としてモニタ504において表示される。ステップ529では、選択された単語が下の階層の部分辞書名であるかどうかを判定し、辞書名である場合ステップ523へ、そうでない場合ステップ530へ進む。ここでは単語「施設」は図2の単語辞書に示されるように下の階層の部分辞書eの辞書名となっているので、ステップ523に戻る。

【0010】ステップ523では、CPUが外部記憶装置503から新たに部分辞書eをメモリに読み込む。このように、処理を繰り返すたび階層が進む。選択された単語の下に部分辞書がなくなると、ステップ530で、認識された単語は目的語となり出力される。ここでは「桜木町」が出力される。

課題

【発明が解決しようとする課題】上記のように従来の音声入力装置においては、一つの単語を伝えるのに、その単語の属する最下層までの部分辞書名を階層毎に入力しなければならない。その階層は単語数が多いほど複雑になるので、大語彙の中からの認識になるほど使用者にとって入力負担が大きいという問題があった。また使用者が最も伝えたいのは目的語であるにも係わらず、目的語の入力が最後で感覚的に違和感を抱かせ、使い勝手がよくないという問題があった。本発明は、上記従来の問題点に鑑み、少ない入力回数で正確な音声入力を実現することを目的としている。

手段

【課題を解決するための手段】発話音声情報を情報化して入力する音声入力手段と、複数の単語を含む部分辞書からなるとともに階層構造を持った単語辞書と、使用する部分辞書を決定する部分辞書決定手段と前記音声入力手段からの入力音声と決定された部分辞書内の単語との一致度を演算する演算手段と、演算された単語のうち最も一致度の高い単語を選択して出力する単語選択手段とを有し、前記部分辞書決定手段は、初めに所定の最下層の部分辞書を決定し、階層変更指令を受けた場合に階層変更をして部分辞書を決定するものとした。

【0013】前記部分辞書決定手段は、上位階層への変更指示を受けた場合、上位階層の部分辞書を決定し、入力により単語選択手段が選択した単語が示す階層に変更し、該階層下の所定の最下層部分辞書を決定することができる。前記部分辞書決定手段により初めに決定される部分辞書は前回使用された最下層の部分辞書であることが望ましい。また前回使用された最下層の部分辞書の代わりに使用頻度の高い最下層の部分辞書を用いても可能である。

【0014】音声入力装置には表示手段を接続し、単語選択手段が選択した単語及び上位の階層構造を表示するのが望ましい。前記表示手段はタッチパネルを合わせ持ち、使用者によってタッチされた階層に対応した階層変更指令を前記部分辞書決定部に出力することも可能である。また、タッチパネルに対するタッチ入力があったときに、変更可能な階層を表示手段上に表示するのが望ましい。

【0015】前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、上位階層への変更指示を受けたとし、設定されている最下層部分辞書より上位の全ての部分辞書を決定し、単語選択手段が選択した単語が示す階層に階層変更を行なうことができる。

【0016】前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、設定されている最下層部分辞書及びそれより上位の全ての部分辞書を決定し、単語選択手段が選択した単語が上位部分辞書内の単語であれば、上位階層への変更指示を受けたとし、前記単語が示す階層に階層変更を行なうこともできる。

【0017】また、最初に入力された音声を記憶する記憶手段を合わせ持ち、階層が変更されたとき、前記単語選択手段は前記記憶手段に記憶された音声を用いて、決定された最下層の部分辞書内の単語との一致度を演算し、出力することができる。

作用

【作用】階層化した単語辞書を使用し、部分辞書を設定する際に、初めに所定の最下層の部分辞書を設定するので、大語彙認識ができるとともに、1回目の単語を目的語とすることもできる。そして一回の入力で認識できなかった場合、階層変更指令で階層変更を行なうので、例えば現在の部分辞書と目的語のある部分辞書と共通する上位階層に変更して最下層の部分辞書を決定する段取りをとることができ、最上位の部分辞書からの入力が必要でなくなり、効率のよい入力が行われる。また、最初から目的語で入力し、辞書の階層変更が認識できなかったときに行なうので、感覚に合致した音声入力が行なえる。

【0019】部分辞書決定手段は、上位階層への変更指示を受けた場合、上位階層の部分辞書を決定し、入力により単語選択手段が選択した単語が示す階層に変更し、該階層下の所定の最下層部分辞書を決定するようにした場合、階層変更があったも、目的語の部分辞書と上位階層との間の部分辞書に対する認識が不要で、階層が上位に変更されても、すぐ目的語の入力となるので、装置使用者を焦らせることなく、入力できる効果が得られる。また部分辞書決定手段が初めに決定する部分辞書は前回使用された最下層の部分辞書となれば、辞書構成を理解し、階層変更が円滑に行なえる。さらに前回使用された最下層の部分辞書の代わりに使用頻度の高い最下層の部分辞書を用いても同じ効果が得られる。

【0020】音声入力装置に表示手段を接続し、単語選択手段が選択した単語及び上位の階層構造を表示手段で表示するようにすると、階層変更が必要のときに、表示された階層構造を頼りに階層変更ができるので、入力負担が軽減される。表示手段はタッチパネルを合わせ持ち、使用者によってタッチされた階層に対応した階層変更指令を前記部分辞書決定部に出力するようにすると、対話式な入力となり一層の入力負担軽減が図られる。また、タッチ入力があったときに、音声入力によって変更可能な階層を前記表示手段上で表示した場合、辞書の構造を表示画面から知ることができ、辞書の構成などの予備知識が無くても階層変更ができる効果が得られる。

【0021】前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、上位階層への変更指示を受けたとし、設定されている最下層部分辞書より上位の全ての部分辞書を決定し、単語選択手段が選択した単語が示す階層に階層変更を行なうようにすると、上位階層の単語のみで、階層変更ができ入力負担がさらに軽減される。

【0022】前記部分辞書決定手段は、目的語に続き入力が行なわれたときに、設定されている最下層部分辞書及びそれより上位の全ての部分辞書を決定し、単語選択手段が選択した単語が上位部分辞書内の単語であれば、上位階層への変更指示を受けたとし、前記単

語が示す階層に階層変更を行なった場合、部分辞書が正確でありながら、単語の誤認識による階層変更が防止される。

【0023】最初に入力された音声記憶する記憶手段を合わせ持ち、階層が変更されたとき、前記単語選択手段は前記記憶手段に記憶された音声を用いて、決定された最下層の部分辞書内の単語との一致度を演算し、出力することで、階層を変更し、最下層の部分辞書が変わっても、目的語の音声を二度と入力する必要がなく、音声の入力回数がさらに減らされる。

【0024】

【発明の実施の形態】次に、本発明の実施の形態を実施例により説明する。図1は、本発明の第1の実施例を構成を示すブロック図である。マイク500により取り込まれた音声信号はA/Dコンバータ501でデジタル情報に変えられて信号処理装置5に入力される。スイッチ505は音声入力用に装置使用者が発話する直前に操作され、信号処理装置5に発話音声の検出タイミングを与えている。外部記憶装置503には単語辞書を記憶させてある。モニタ60はタッチパネル付きモニタで、信号処理装置5の処理結果を表示するとともに、タッチパネルで部分辞書の変更が行なえるようになっている。

【0025】単語辞書は図2に示すような階層化した単語辞書が用いられる。図において部分辞書間の連線が上下層間の繋がり関係を表示し、枠内に突出したのは下の階層を代表する単語についての表示である。「住所」と「施設」の二つの単語で構成する部分辞書aは単語辞書の最上層にあり、各単語に下位階層辞書として「住所」からは政令指定都市に従った分類で都道府県名を記した部分辞書bを設け、各都道府県について、市区名を記した部分辞書cと区村名を記した部分辞書dが順次に設けられている。

【0026】部分辞書aの単語「施設」の下位には「駅」、「デパート」、「ホテル」のような「施設」の種類名を記した単語の部分辞書eが続いている。「駅」の下位に都道府県名を記した部分辞書fが設けられている。部分辞書fからは交通会社名を記した部分辞書gと、部分辞書gに繋がっている駅名の部分辞書hが都道府県毎に設けられている。

【0027】また、単語「デパート」も都道府県名を記した下層の部分辞書iと、市名を記した部分辞書jと、「デパート」名を記した部分辞書kを持ち、都道府県別に市区の部分辞書と区村の部分辞書に繋がっている。部分辞書d、部分辞書h、部分辞書kは最下層部分辞書であり、目的語が登録され、認識したい単語はここで照合され認識される。

【0028】信号処理装置5は、音声入力部53においてスイッチ505が操作されるまでマイク500からデジタル化された音声信号の平均パワーを演算している。その平均パワー値より瞬間パワーが大きくなったときに、発話が開始したと判断し、音声信号の取り込みを開始する。その音声信号は単語選択部52において単語認識される。この際最初に使われる単語辞書は部分辞書決定部51が外部記憶装置503から読み込んだ最下位階層にある部分辞書である。

【0029】単語出力部54はその認識の結果と単語の上位階層の名称をモニタ60に一

時出力するが、装置としての出力は保留する。この間モニタ60上に表示された上位階層の名称をタッチ入力すれば部分辞書の変更ができる。所定時間内にタッチパネルへのタッチが無ければ、単語出力部54は認識の結果を装置の出力として出力する。

【0030】部分辞書の変更があった場合は、部分辞書決定部51はタッチされた名称の部分辞書を外部記憶装置503から読み込み、装置使用者はスイッチ505を操作して下位辞書を決定するための発話をする。部分辞書決定部51はそれに係わる所定の最下位部分辞書を決定し、外部記憶装置503から読み込む。その後再び音声入力をし単語選択部52において単語認識される。最初に入力するあるいは階層変更後に入力される最下層の部分辞書は前回使用した部分辞書あるいは使用頻度の最も高い部分辞書である。

【0031】次に、図3のフローチャートに従って装置の作動の流れを説明する。まず、ステップ101において、音声入力の初期状態で、階層の最下位の部分辞書を最初の認識対象として部分辞書決定部51において設定する。初期設定として最も使用頻度の高い部分辞書あるいは前回使用された最下層の部分辞書が用いられるが、ここでは例えば図2中の部分辞書hが設定されたとする。

【0032】装置使用者はスイッチ505を操作して、マイク500に発話をする。ステップ102では、音声入力部53がスイッチ505が操作されたか否かを検出し、操作された場合、マイク500が動作してからの音声の平均パワー値を算出してステップ103へ進む。ステップ103では、決定された部分辞書を前記外部記憶装置503から部分辞書決定部51に読み込む。ここでは部分辞書hが読み込まれる。

【0033】ステップ104では、使用者の発話音声取り込みを開始する。すなわち入力した音声信号を絶えず算出した音声の平均パワーと比べ、瞬間パワーが所定値以上大きくなったとき、使用者の発話開始と判断し、発話音声の取り込みを開始する。ここでは、例えば使用者は目的語として「桜木町」を発話したとする。ステップ105では、部分辞書決定部51に読み込まれた部分辞書h内の単語と発話音声との一致度を単語選択部52において演算する。なお、ここでの処理は音声信号をとりながら所定時間で区切った音声区間部分と各単語との比較が行なわれており、音声信号の取り込みはそれに同時に進行されている。

【0034】ステップ106で、音声信号の瞬間パワーが平均パワー以下になったとき、使用者が発話完了と判断し、音声取り込み終了してステップ107へ進む。ステップ107では、単語選択部52で演算された一致度の最も高い単語を認識の結果として単語出力部54に出力する。

【0035】ステップ108では、単語出力部54はその認識の結果と単語の上位階層の名称をモニタ60に一時出力し表示させるが、装置としての出力は保留する。図4はモニタ60の表示画面である。hが選択された単語の辞書で、a、e、f、gはその単語の上位辞書である。hの枠から表示されているのは目的語であり、その上の枠に並べられているのが目的語に関連する上位辞書内の単語で、右から左へ階層順位の増加を表示している。

階層を変更する場合に、単語が映っているところをタッチすれば、階層の変更ができる。そしてステップ109で、所定時間内にタッチパネルへのタッチ入力が無ければ、ステップ110において「桜木町」という単語の認識結果を装置の出力として出力する。この場合、階層を変更する必要がなく、発話回数が一回で終了する。

【0036】もし、使用者がステップ104で「そごう」と発話していたとすれば、初期の部分辞書がhであったため、他の単語が表示されてしまう。その場合、上位の階層を変更する必要が生じるため、使用者は図4中のeをタッチ入力する。

【0037】そのタッチ入力がステップ109で検出されると処理がステップ111へ進む。ステップ111では、タッチされた階層をどの階層に変更可能かモニタ60に表示する。ここでは「駅」、「デパート」、「ホテル」を表示する。図5はその表示様子を示している。ステップ112では、部分辞書決定部51が新たに部分辞書eを外部記憶装置503から読み込む。

【0038】ステップ113では、タッチして表示された上位階層の中から選択して音声入力する。「そごう」の場合は使用者はそれに関連のある「デパート」を音声で入力し、単語選択部52で認識される。本ステップはステップ102からステップ107までと同様の処理を行なってもよい。ステップ114では、部分辞書決定部51が、認識された単語「デパート」に繋がる最下層の部分辞書kを決定する。

【0039】ステップ115では、再度音声入力やり直しを使用者に報知するため、モニタ60に再度入力の指示マークを表示して、ステップ102に戻る。その後使用者が「そごう」を発話してステップ102からステップ108までの処理が再び行なわれて、ステップ110において、「そごう」という単語が出力される。

【0040】マイク500、A/Dコンバータ501、音声入力部53は音声入力手段を構成している。部分辞書決定部51は辞書決定手段を構成している。単語選択部52は単語選択手段を構成している。モニタ60は表示手段を構成している。

【0041】本実施例は以上のように構成され、第1の音声の入力が行なわれた場合、音声信号と所定の最下層の部分辞書の単語との一致度が演算されるので、第1の単語を目的語とすることができ、使用者の入力負担が軽減されることになる。また、部分辞書の変更が必要なとき、上位の階層の下に複数の階層が存在しても、その最下層の部分辞書を設定して使用するので、階層変更による入力負担増加が少ない。そして、最初に使用する部分辞書を使用頻度の高いものあるいは前回使用したものとしているので、使用者は出力結果を容易に理解し、階層変更が円滑に行なえる。

【0042】また、モニタが単語を表示するときに単語の上位階層構造を表示するので、どの階層を変更したいかを容易に判断できる。タッチ入力があったとき、音声入力によってどのような階層に変更可能かを表示することで、辞書の構成を知らなくても、どの階層に変更するかの操作ができる。

【0043】図6は、第2の実施例の構成を示すブロック図である。この実施例は、図1

に示す第1の実施例における信号処理装置5の代わりに音声記憶部66を設けた信号処理装置6を用いる。音声信号記憶部66は音声入力部63の音声信号を単語選択部62に出力するとともに、一回目の音声信号は音声記憶部66が記憶する。部分辞書変更があった場合、音声記憶部66に記憶された音声信号を用いて部分辞書との一致度を判定し、目的語を認識する。その他は第1の実施例と同様である。

【0044】次に、図7のフローチャートに従って装置の作動の流れを説明する。ステップ201～ステップ215まではステップ207を除いて第1の実施例における図3のフローチャートのステップ101～ステップ114と同様の処理を行なう。すなわちまず、階層の最下位の部分辞書を最初の認識対象として設定し、外部記憶装置503から読み込む。使用者の発話音声信号を取り込みながら設定された部分辞書内の単語と一致度を演算する。そして音声信号の取り込みが終了すると、ステップ207において、音声記憶部66が音声信号を記憶する。一致度の最も高い単語が認識の結果として選択され、上位階層を含めてモニタ60において表示される。

【0045】そして、所定時間内にモニタ60のタッチパネルに対するタッチ入力が無ければ、選択された単語を装置の出力として出力するが、タッチ入力があった場合、タッチされた階層に部分辞書の変更を行なう。階層変更された部分辞書が新たに外部記憶装置503から読み込まれる。その後階層変更のための発話をし、音声入力によって階層を示す単語が特定されると、その階層下の最下位部分辞書を外部記憶装置503から読み込む。

【0046】その後、第1の実施例では再度音声入力やり直しを使用者に報知するが、本実施例では、ステップ216において、音声記憶部66から記憶した音声信号を用い、最下層の部分辞書内の単語との一致度を演算する。各単語から一致度の最も高い単語がステップ208において選択される。そしてステップ210にモニタ60へのタッチ入力がないと判定されると、ステップ211で認識した単語を装置の出力として出力する。

【0047】本実施例によっても、第1の実施例と同様の効果が得られるとともに、部分辞書の変更後、再度最初に目的語を発話し音声入力する必要がないので、入力回数が減少されることになる。

【0048】次に、第3の実施例について説明する。図8は、第3の実施例の構成を示すブロック図である。この実施例は、図6に示す第2の実施例における信号処理装置6の代わりに信号処理装置7を用いる。音声信号記憶部66は音声入力部73の音声信号を単語選択部62に出力するとともに、一回目の音声信号は音声記憶部66が記憶する。部分辞書変更がある場合、音声記憶部66に記憶された音声信号を用いて認識する。

【0049】第1あるいは第2の実施例では階層の変更はモニタ60のタッチパネルで行なったが、本実施例は音声による階層変更を行なう。すなわち第一回目の音声入力を目的語に対応させ、以後の音声入力は階層変更に対応させている点が第1あるいは第2の実施例と異なる。モニタ60の代わりに表示のみのモニタ70を使用する。その他は第1の実施例と同様である。

【0050】次に、図9のフローチャートに従って装置の作動の流れを説明する。ステップ201～ステップ216まではステップ310、ステップ313を除いて第2の実施例における図7のフローチャートと同様の処理を行なう。すなわちまず、階層の最下位の部分辞書を最初の認識対象として設定され、外部記憶装置503から読み込まれる。使用者の発話音声信号を取り込みながら設定された部分辞書内の単語と一致度を演算する。そして音声信号の取り込みが終了すると、音声記憶部66が音声信号を記憶する。そしてステップ208において一致度の最も高い単語が認識の結果として選択され、ステップ209において上位階層を含めてモニタ70において表示されるとステップ310へ進む。

【0051】ステップ310では、音声入力部73がスイッチ505が操作されたか否かを検出し、操作された場合、音声入力があるため、ステップ313へ進む。ステップ313では現在使用中の部分辞書の上位全ての部分辞書を外部記憶装置503から読み込む。その後第1、第2の実施例と同様に上位階層の音声入力がなされ、階層を示す単語が特定されると、その階層下の最下位部分辞書を外部記憶装置503から読み込む。

【0052】ステップ216において、音声記憶部66から記憶した音声信号を用い、部分辞書内の単語との一致度を演算する。各単語から一致度の最も高い単語がステップ208において選択される。ステップ310においてスイッチ505に対する操作がないことが検出されると、ステップ211で認識した単語を装置の出力として出力する。このように、例えば最初に部分辞書hが設定されていて、入力する単語は辞書にない「そごう」の場合、使用者が「デパート」を発話すれば、図10に示すように「そごう」の上位辞書を含むハッチングした部分辞書が自動的に設定されるので、階層変更があっても煩わしい操作はない。

【0053】すなわち、最初の発話に対し、装置が使用者の指定したい階層と異なる階層の単語を出力する場合、使用者が指定したい階層名称を発話するという手順をとることで、発話、返答、詳しい説明という自然な対話を実現でき、最も自然な感覚で音声入力ができる効果が得られる。各階層の部分辞書が設定されるので、音声による認識率の低下が懸念されるが、一般に上位の部分辞書内の単語は、分類のための単語であり、単語数は最下層の部分辞書内の単語数に比較して極めて少ないため、認識率は十分に使用に堪える。

【0054】次に、第3の実施例の変形例について説明する。前記各実施例では、最下層の部分辞書が正確に設定されても、騒音などの影響で単語を誤認識することがある。この場合すべての上位階層の部分辞書を設定し直しても辿り着いたのはもとの部分辞書であり、余計な入力回数を作る結果となる。この変形例はそれに対処するためのもので、誤認識があっても、同じ部分辞書内の単語なら、階層変更せずに処理できるようにしている。

【0055】図11は装置作動の流れを示すフローチャートである。このフローチャートはステップ412、ステップ413、ステップ414を図9の第3の実施例におけるステップ313、ステップ214に置き換えて構成される。すなわちまず、階層の最下位の部分辞書を最初の認識対象として設定され、外部記憶装置503から読み込まれる。使用者

の発話音声信号を取りながら設定された部分辞書内の単語と一致度を演算する。そして音声信号の取り込みが終了すると、音声記憶部66が音声信号を記憶する。一致度の最も高い単語が認識の結果として選択され、ステップ209において上位階層を含めてモニター70において表示される。

【0056】そして、ステップ310において、音声入力部73がスイッチ505が操作されたか否かを検出し、操作された場合、音声入力があるため、ステップ412へ進む。ステップ412では、現在使用中の部分辞書およびその上位全ての部分辞書を読み込む。ステップ413では、使用者が発話した音声を入力し、読み込まれた部分辞書内の単語との一致度を演算する。

【0057】ステップ414では一致度の最も高い単語が上位階層の部分辞書内にあるか否かを判定する。単語が上位階層の部分辞書内にある場合、階層変更するためのものであり、ステップ215へ進み、第3の実施例と同様の処理を行なう。単語が最下層の部分辞書内にある場合、既に認識された目的語であるためステップ209へ進み、ステップ310でスイッチに対する新たな操作が行なわれていないかの判定を経てステップ211において単語が出力される。

【0058】このように、もし最初に部分辞書hが設定されていて、入力したい単語は「桜木町」の場合、騒音などによる「桜木町」の音声信号に歪みが生じ、「関内」と誤って表示される可能性がある。また「桜木町」を発話するつもりで「関内」と誤って発話することが考えられる。このような部分辞書が正しい場合でも、図12にハッチングで示すように部分辞書hを含む全ての上位階層の部分辞書が自動的に設定されるので、「桜木町」と発話し直しすれば、階層変更をせずに音声入力ができる。これによって階層の変更は必要などきのみ行なわれることになり、音声入力負担がさらに軽減される。このほか第3の実施例のように発話、返答、詳しい説明という自然な対話形式で音声入力ができる効果も得られる。

効果

【発明の効果】以上説明したように、本発明によれば、1回目の単語を目的語とすることができるので、まず目的語で入力し、正確でなかったら階層を変更して目的語を再度入力して認識し、自然な感覚で音声入力ができる。また辞書を変更する際、例えば現在の部分辞書から遡り、目的語と共通する内容の上位部分辞書に変更して最下層の部分辞書を決定する段取りをとるので、目的語が認識されるまでの入力回数が少なく、入力負担が軽減される【0060】部分辞書決定手段は、上位階層への変更指示を受けた場合、上位階層の部分辞書を決定し、音声入力により単語選択手段が選択した単語が示す階層に変更し、該階層下の所定の最下層部分辞書を決定するようにした場合、階層変更があったも、目的語の入っている部分辞書に至るまでの部分辞書に対する認識が不要で、大きな入力負担とな

らない効果が得られる。また部分辞書決定手段が初めに決定する部分辞書は前回使用された最下層の部分辞書となれば、辞書構成を理解し、階層変更が円滑に行なえる。さらに前回使用された最下層の部分辞書の代わりに使用頻度の高い最下層の部分辞書を用いても同じ効果が得られる。

【0061】音声入力装置に表示手段を接続し、単語選択手段が選択した単語及び上位の階層構造を表示手段で表示するようにすると、階層変更が必要のときに、表示された階層構造を頼りに階層変更ができるので、入力負担がさらに軽減される。前記表示手段はタッチパネルを合わせ持ち、使用者によってタッチされた階層に対応した階層変更指令を前記部分辞書決定部に出力するようにすると、対話式な入力となり一層の入力負担軽減が図られる。また、タッチ入力があつたときに、音声入力によって変更可能な階層を前記表示手段上で表示した場合、辞書の構造を表示画面から知ることができ辞書の構成に対する予備知識が無くても階層変更ができる効果が得られる。

【0062】前記部分辞書決定手段は、目的語に続き音声入力が行なわれたときに、上位階層への変更指示を受けたとし、設定されている最下層部分辞書より上位の全ての部分辞書を決定し、単語選択手段が選択した単語が示す階層に階層変更を行なうようにすると、階層変更に対する操作が不要で入力負担が一層軽減される。

【0063】前記部分辞書決定手段は、目的語に続き音声入力が行なわれたときに、設定されている最下層部分辞書及びそれより上位の全ての部分辞書を決定し、単語選択手段が選択した単語が上位部分辞書内の単語であれば、上位階層への変更指示を受けたとし、前記単語が示す階層に階層変更を行なった場合、部分辞書が正確でありながら、単語の誤認識による階層変更が防止され、無駄な入力を生じさせない効果が得られる。

【0064】最初に入力された音声記憶する記憶手段を合わせ持ち、階層が変更されたとき、前記単語選択手段は前記記憶手段に記憶された音声を用いて、決定された最下層の部分辞書内の単語との一致度を演算し、出力することで、階層を変更し、最下層の部分辞書が変わっても、目的語の音声入力を二度と行なう必要がなく、音声の入力回数がさらに減らされる。また、最初の発話に対し、装置が使用者の指定したい階層と異なる階層の単語を出力する場合、第2の実施例と同様に使用者が指定したい階層名称を発話するという手順をとることで、発話、返答、詳しい説明という自然な対話を実現でき、最も自然な感覚で音声入力ができる効果が得られる。

図の説明

【図面の簡単な説明】

【図1】第1の実施例の構成を示すブロック図である。

【図2】階層化した単語辞書の構成を示す図である。

【図3】第1の実施例のフローチャートである。

【図4】 モニタ表示画面を示すブロック図である。

【図5】 部分辞書を変更時のモニタ表示画面を示す図である。

【図6】 第2の実施例の構成を示すブロック図である。

【図7】 第2の実施例のフローチャートである。

【図8】 第3の実施例の構成を示すブロック図である。

【図9】 第3の実施例のフローチャートである。

【図10】 最下位部分辞書を変更時に仮設定された上位部分辞書を示す図である。

【図11】 第3の実施例の変形例を示すフローチャートである。

【図12】 最下位部分辞書を変更時に仮設定された部分辞書を示す図である。

【図13】 従来例の構成を示すブロック図である。

【図14】 従来例のフローチャートである。

【符号の説明】

5、6、7 信号処理装置

51、71 部分辞書決定部

52、62 単語選択部

53、63、73 音声入力部

54、64 単語出力部

60、70、504 モニタ

66 音声記憶部

500 マイク

501 A/Dコンバータ

502 信号処理装置

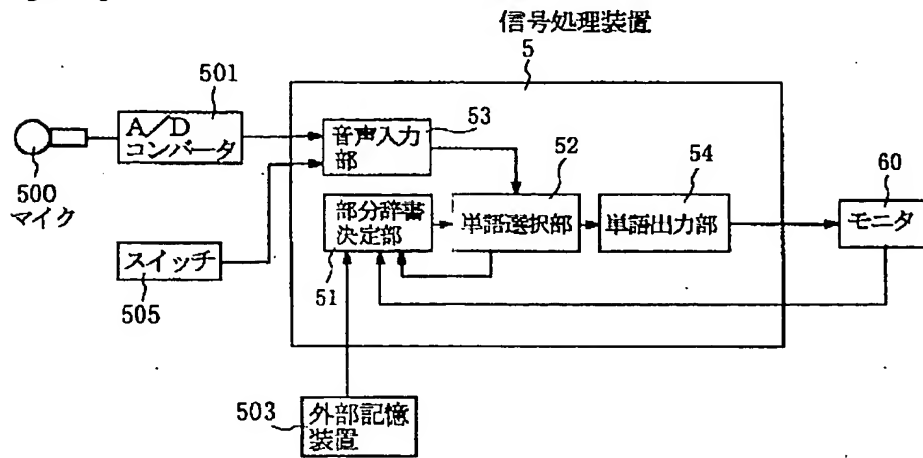
503 外部記憶装置

505 スイッチ

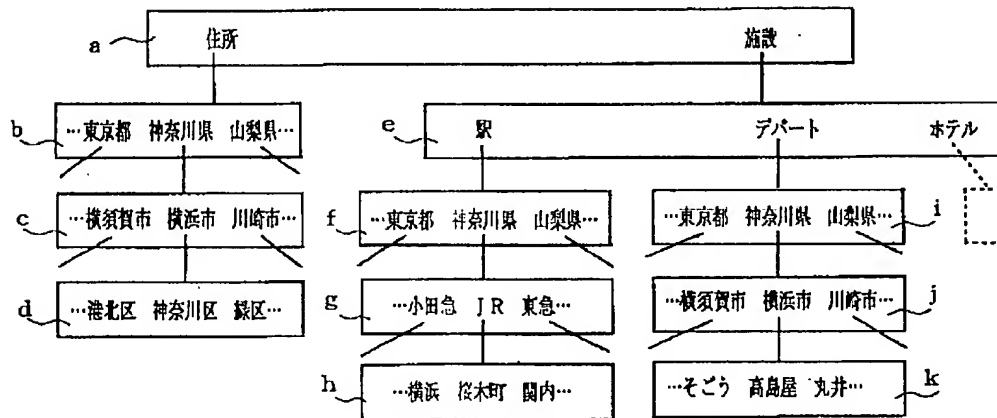
a、b、c、d、e、f 部分辞書

g、h、i、j、k 部分辞書

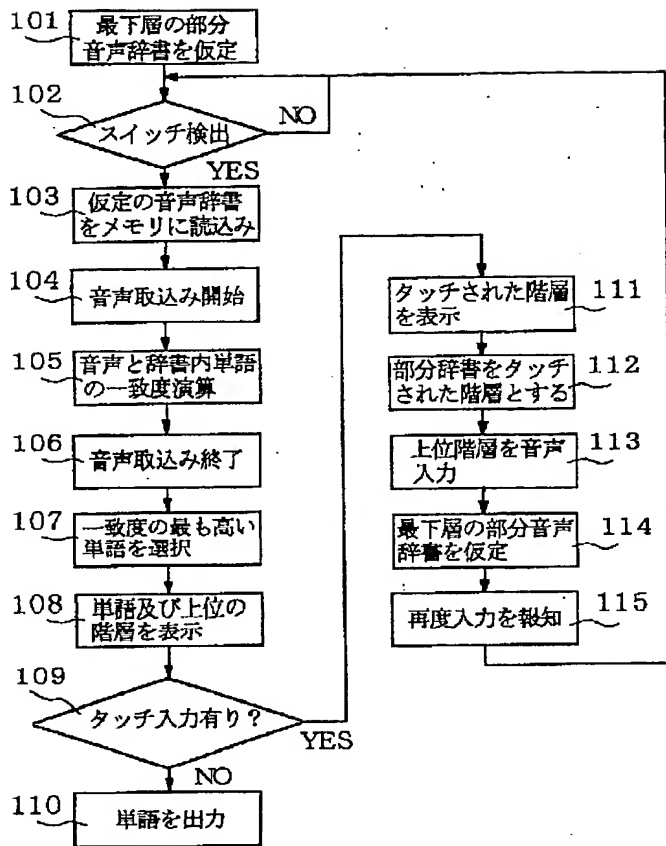
【図1】



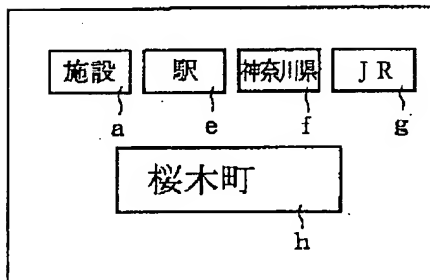
【図2】



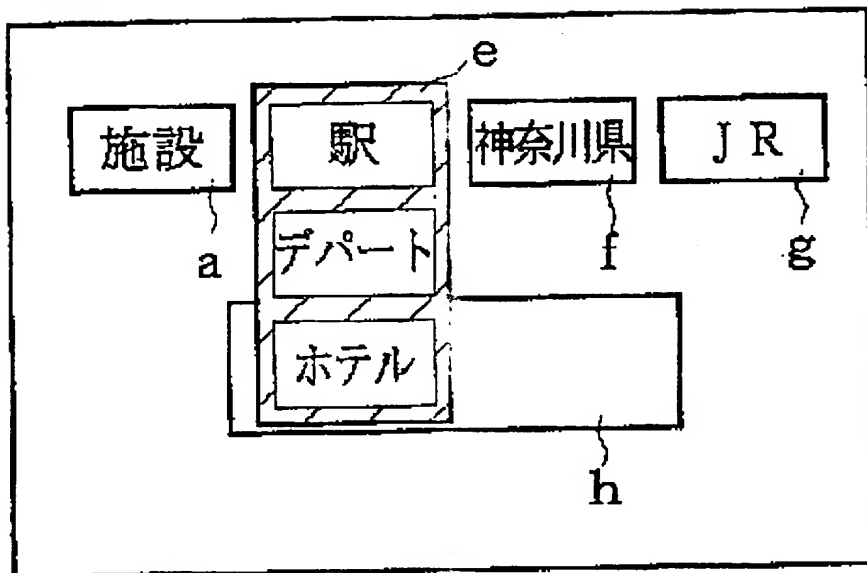
【図3】



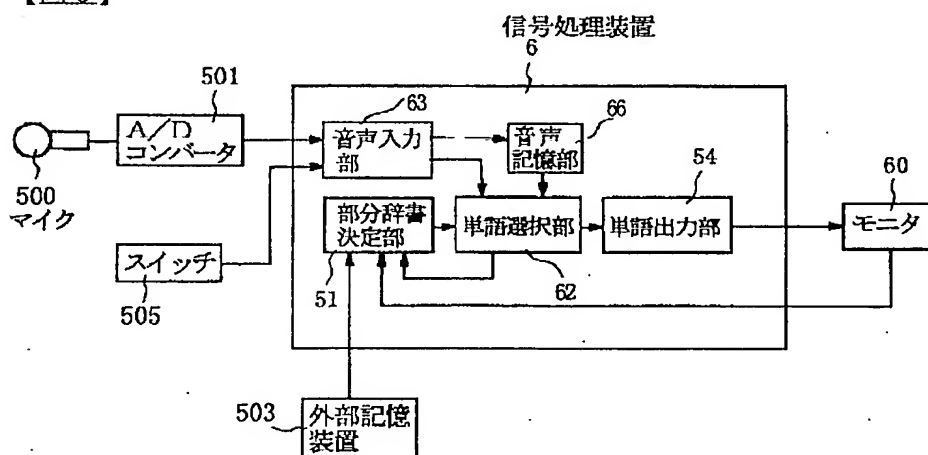
【図4】



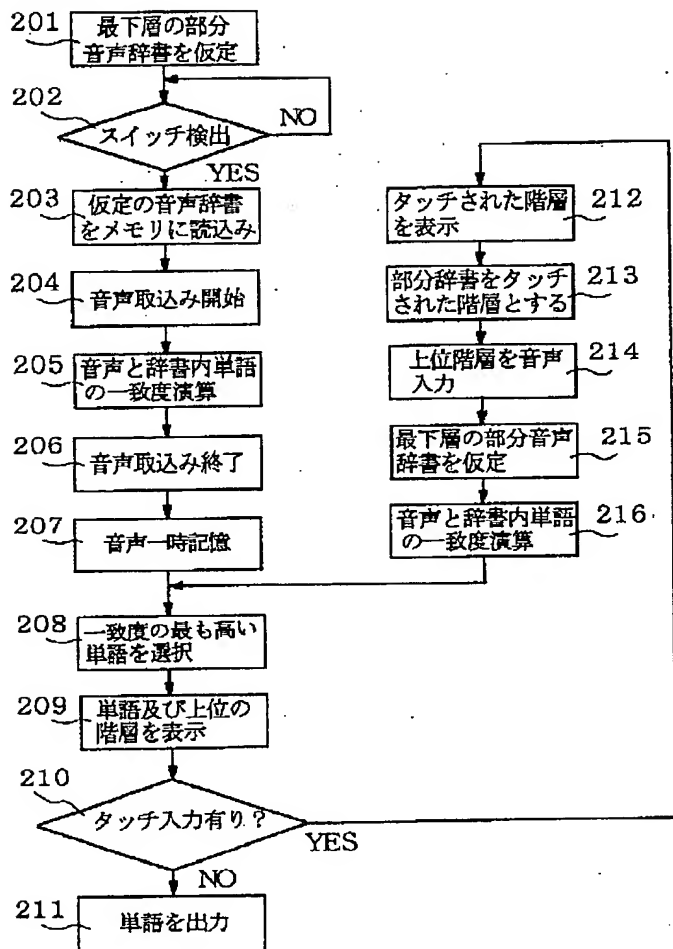
【図5】



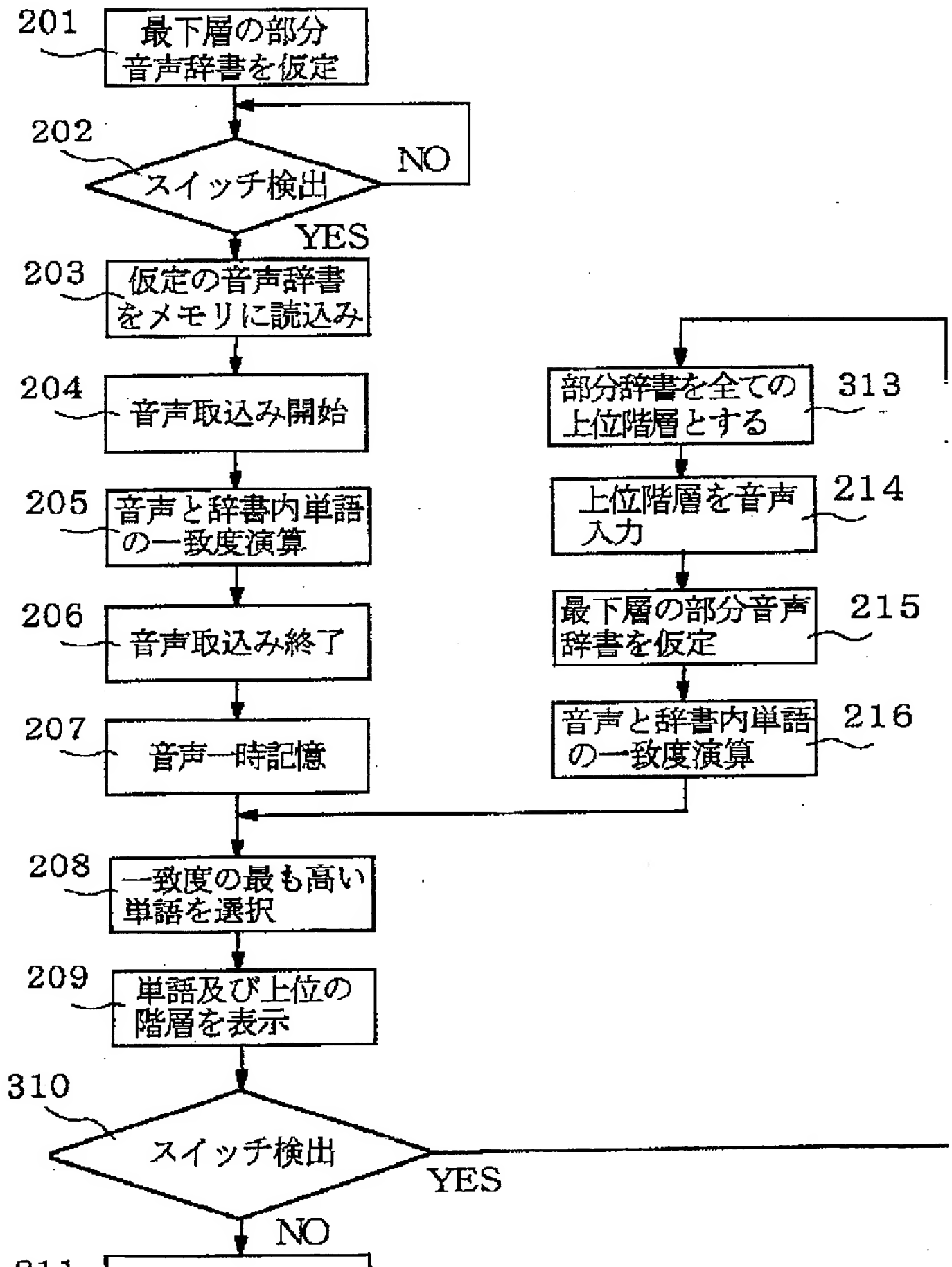
【図6】



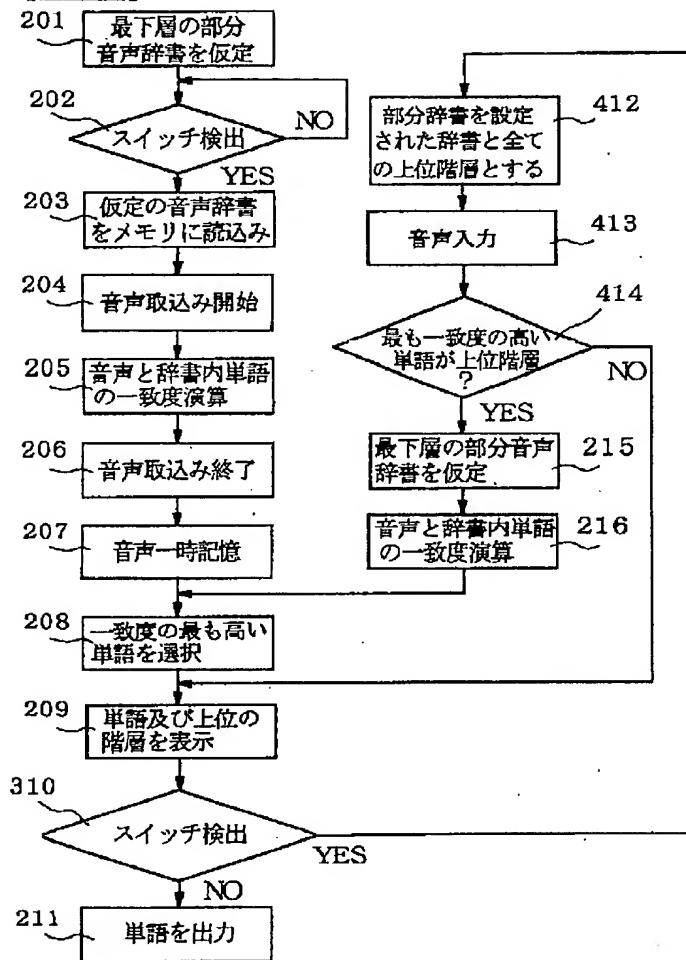
【図7】



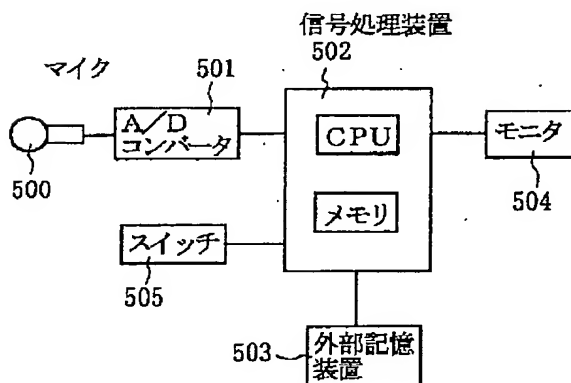
【图9】



【図1.1】



【図1.3】



【図14】

